

Success and Failure in Skill Qualification
Testing: Troop Views

Joan Harman, Ph.D.
Research Psychologist
US Army Research Institute
for the Behavioral and Social Sciences

A contribution to the ongoing development of the SQT system would be made by exploring reasons for success and failure in testing. One source of this information is enlisted soldiers who take the tests. Accordingly, enlisted soldiers drawn from combat arms, combat support and combat service support MOSs were presented with questions taken from 1980 written SQTs. The questions used were those on which substantial numbers of testees had succeeded or failed. The soldiers were asked to account for the widespread success or failure by other members of their MOS. Most attributed both success and failure to whether tasks being tested were performed as part of unit duties. The lowest number of attributions concerned personal efforts to prepare for written testing. This suggests a need to organize and standardize written test training and either to eliminate or clearly stipulate individual soldiers' responsibilities for test preparation.

Success and Failure in Skill Qualification Testing: Troop Views

Introduction

Since the first Skill Qualification Tests were fielded in 1977, the system has been undergoing continuing development and revision to make it responsive to Army needs. This demands increasing efforts to make the system a valid reflection of soldiers' skills and a clear indicator to managers and planners of individual readiness. Accordingly, all facets of the Army have been concerned and involved with Skill Qualification Testing.

In 1980 the Office of the Undersecretary of the Army expressed a special interest in Skill Qualification Testing, with particular emphasis on interpreting low test results. At least in part, this added interest was prompted by media reports about soldiers performing poorly on these tests. One approach to gathering this kind of information is the subject of the research reported in this paper.

Of the three components of SQT -- on-the-job testing, hands-on testing and written testing -- failures have occurred most commonly on the written component. In fact, soldiers have a history of high levels of success on the other two components. Therefore, this paper deals exclusively with performance on the SQT's written test, called the Skill Component (SC).

Of sources of information that would account for soldiers' poor test performance, the most obvious may be test developers, commanders and trainers. Test developers, however, tend to emphasize problems in training programs whereas training managers and trainers focus on problems in the tests. Both of these viewpoints must be carefully weighed, since it is unlikely that any program for training or testing has no margin at all for improvement. There is another source of information about test performance, however, and it is one that may reflect no bias in behalf of either training or testing. This source is the enlisted soldiers who take SQTs. Insofar as individual soldiers are able to identify causes for their own or their unit's failure to answer questions about critical tasks correctly, they could provide clues to identifying training and/or testing variables which would enhance development of the Skill Qualification Testing system. An additional consideration is that even though the main emphasis here is on unsuccessful performance, reasons for success in testing could yield valuable information about reasons for failure.

To explore this hypothesis, enlisted soldiers assigned to Combat Arms, Combat Support and Combat Service Support MOS who took the 1980 Skill Component were interviewed. They were asked to account for widespread success on some SQT questions and for widespread failure on others.

METHOD

Subjects

Ninety-four enlisted soldiers participated in this research. They were drawn from two Combat Arms, one Combat Support, and two Combat Service Support MOS and stationed at Ft. Bragg, Ft. Carson, and Ft. Hood.

Procedure

Skill Component questions were selected from 1980 SQT Item Analyses supplied to ARI by the Army Training Support Center. Criteria for question selection were:

- 1 - the set of questions for each MOS included both common soldier and MOS specific tasks,
- 2 - each question had only one correct answer,
- 3 - more than 50% of the soldiers tested answered the questions correctly/incorrectly. (The percent GO range for successful questions is 70-84 with a median of 81 and for unsuccessful questions is 12-48 with a median of 29.)

For each MOS, one successful and five unsuccessful questions were typed on 5x8 index cards in the same format in which they appeared in the 1980 SQT. That is, each card displayed one task title and number, one general situation, one question and its accompanying answer choices. Appropriate cards were presented one at a time to participants. Each soldier was asked to account for the wide ranging success or failure of soldiers in the MOS on that question. After the soldier's response was recorded, the answer selected was either confirmed as correct or was corrected. This procedure was repeated in the same way for all six questions for each MOS.

RESULTS

Table 1 shows percent of soldier's responses, ranked from highest to lowest frequencies, in four general categories: performance attributed to events or practices occurring in units, performance attributed to training, performance attributed to test content, and performance attributed to soldiers' personal characteristics. It's important to note that subcategories under each main heading are not intended to be mutually exclusive. For example, under the category Unit Factors, it seems reasonable to assume that a task with which a soldier has had no experience is one that is not job related. Also, a Specialty Task, one that is performed by only a limited number of MOS members, would also be unrelated to the jobs of other members. Although somewhat interrelated, these subcategories were established because every effort was made to reflect the actual responses soldiers made and to avoid interpreting them. These data are summarized by MOS type in Table 2.

TABLE 1

PERCENT RESPONSES BY 94 ENLISTED SOLDIERS ACCOUNTING FOR
SUCCESS AND FAILURE ON SKILL COMPONENT QUESTIONS

LESS THAN 50% OF POPULATION ANSWERED QUESTIONS CORRECTLY
(470 TOTAL RESPONSES)

	11H	13B	12C	71L	76X	
UNIT FACTORS						
PROCEDURE DIFFERENCES	10	8	19	< 1	10	
NO JOB RELEVANCE	1	8	6	24	10	
NO EXPERIENCE	2	5	14	9		
SPECIALTY TASK	4	8	10	11	5	
LOW PRIORITY TASK		1	1		5	
WORK OUT OF MOS			1		5	
TOTAL						34
TRAINING FACTORS						
NO TRAINING	14	9	14	7	35	
NONE SINCE BASIC		1	3			
NONE SINCE AIT	2			3		
NONE SINCE SQT	4		1		10	
INFREQUENT	9	2				
INCOMPLETE	5	< 1	1	< 1	5	
TOTAL						19
TEST FACTORS						
CONFUSING SITUATION/ QUESTION/ANSWER	8	11	1	9		
POOR QUESTION	2	3		< 1		
ANSWER WRONG	6	7	1		1	
EQUIPMENT OBSOLETE	4		1			
TOTAL						15
SOLDIER FACTORS						
UNLEARNED/FORGOTTEN DETAILS	8	1	1	9	5	
INCOMPLETE KNOWLEDGE			3	< 1		
NO STUDY	2	5	1			
MISUNDERSTANDING/ MISREADING	2	6		5		
CALCULATION PROBLEMS	9		1			
TOTAL						14

MORE THAN 50% OF POPULATION ANSWERED QUESTION CORRECTLY
(94 TOTAL RESPONSES)

UNIT FACTORS						
PERFORMED FREQUENTLY	15	26		57	25	
FUNDAMENTAL MOS TASK	20	26				
TOTAL						41
TRAINING FACTORS						
TRAINED FREQUENTLY	45	17	50	4		
TOTAL						24
TEST FACTORS						
ANSWER EASY/OBVIOUS		9	14	14	50	
TOTAL						11
SOLDIER FACTORS						
TEST SOPHISTICATION	5	9	7	9		
TOTAL						7

Table 2

Percent Responses by 94 Enlisted Soldiers Accounting for
Success and Failure on Skill Component Questions

	<u>FAILURE</u>				<u>SUCCESS</u>			
	<u>Combat Arms (275 resp)</u>	<u>Combat Support (70 resp)</u>	<u>Combat Service Support (125 resp)</u>	<u>X</u>	<u>Combat Arms (55 resp)</u>	<u>Combat Support (14 resp)</u>	<u>Combat Service Support (25 resp)</u>	<u>X</u>
Unit Factors	25	51	44	40	44	0	52	48
Training Factors	20	20	17	19	27	50	4	27
Test Factors	20	7	9	12	5	7	20	10
Soldier Factors	15	7	15	12	1	7	8	7

The prominent feature of both tables is that the same rankings occur for both successful and unsuccessful performance. Another feature is that enlisted soldiers' attributions diverged from the tendencies of trainers and test developers to point to test and training factors and, instead, emphasized the importance of performing tasks as part of unit duties in order to be prepared to answer questions about them on the SC. The popularity of this response is demonstrated in Table 2. For both kinds of questions, the sums of the mean values of the remaining categories approximate the single values for Unit Factors. A description of Table 1 subcategories follows in order to reflect in greater detail what soldiers may be thinking.

Under the category Unit Factors dealing with unsuccessful performance, the first subcategory, Procedure Differences, encompasses responses indicating that unit procedures don't always match procedures detailed in Soldier's Manuals on which test questions are based. For example, Bridge Crewmen (12C) were asked about the first thing they would do to construct an individual defensive position in an area designated by their squad leader. The Soldier's Manual answer is "put in sector-of-fire stakes." Most crewmen, however, chose the answer "dig a hasty hole" and insisted that they would follow this procedure regardless of the Soldier's Manual directives. The subcategory No Job Relevance means that soldiers claimed testees chose the wrong answer because the task being tested is not performed as part of unit duties. This was especially true of Administrative Clerks (71L) whose duties tend to reflect the

Under the category Unit Factors dealing with unsuccessful performance, the first subcategory, Procedure Differences, encompasses responses indicating that unit procedures don't always match procedures detailed in Soldier's Manuals on which test questions are based. For example, Bridge Crewmen (12C) were asked about the first thing they would do to construct an individual defensive position in an area designated by their squad leader. The Soldier's Manual answer is "put in sector-of-fire stakes." Most crewmen, however, chose the answer "dig a hasty hole" and insisted that they would follow this procedure regardless of the Soldier's Manual directives. The subcategory No Job Relevance means that soldiers claimed testees chose the wrong answer because the task being tested is not performed as part of unit duties. This was especially true of Administrative Clerks (71L) whose duties tend to reflect the needs of the administrative office to which they are assigned, and who have few or no opportunities to type, for example, non-military letters. No Experience covers responses that indicated soldiers never performed the task. Specialty Task means that task performance is restricted to designated unit members. For example, Administrative Clerks claimed that such a task is Typing Military Orders and Bridge Crewmen pointed to Classifying Vehicles for bridge crossings. Low Priority Tasks tend to be Common Soldier tasks such as first aid and map reading. Soldiers indicated that MOS tasks receive much greater unit emphasis than do Common Soldier tasks. Work Out of MOS is a relatively slender category. It is included, however, because it is a fact of Army life that presents special training and testing difficulties. Very few Subsistence Supply Specialists (76X) in Continental United States perform MOS tasks on a daily basis. Also, even though few Bridge Crewmen attributed SC failure to this problem, an informal interchange with the group prior to interviewing included the fact that substantial numbers of them work out of their MOS.

The subcategories under Training Factors are self explanatory.

Under Test Factors, the first two subcategories, Confusion about test materials and the opinion that test materials are Poor could be related. The difference between a confusing question and a poorly constructed question may be marginal. Nonetheless, efforts to avoid subjective interpretation establishes these are separate categories. The Answer Wrong subcategory covers flat statements that the answer soldiers were informed was the correct one is, in fact, the wrong answer to the question. In some cases, soldiers declared that all response choices were wrong. Equipment Obsolete refers to responses in which it was claimed that SC questions dealt with outdated equipment. For example, Antiarmor Crewmen (11H) questioned about engaging targets with an M72A2 LAW reported that they don't use the LAW because it has been supplanted by the TOW.

Under Soldier Factors, the subcategory Unlearned/Forgotten Details most often included details involving numbers; for example, the ratio of chest compressions to lung inflations in cardiopulmonary resuscitation.

Attributions about successful SC performance are virtually a mirror image of those dealing with SC failure. This observation is supported by the fact that analysis of differences between mean values of both sets of questions (see Table 2) using the Randomization Test for Matched Pairs shows no significance ($p = .05$). The only subcategory that may not be self explanatory is Test Sophistication under Soldier Factors. It covers responses indicated that soldiers are testwise. For example, those who reported that they didn't know or weren't sure of the correct answer but were able to eliminate clearly incorrect choices or picked up clues to the correct answer in the General Situation or in the test question.

A final significant feature of the Tables is that in no case do percentages sum to 100. Missing values can all be categorized as "I don't know" responses.

DISCUSSION

The opinions of enlisted soldiers asked to account for widespread failure of individuals in their MOS on SC questions stress the influence of unit activities on test results as opposed to training, test or personal characteristics. This can be interpreted to mean that soldiers don't expect to be prepared to answer test questions on critical tasks that are not part of unit duties. That is, they appear to discount the influence of training on SC success and they show little acceptance of personal responsibility for test preparation. This seems to indicate a need for clearer communication about expectations surrounding SC preparation.

One area of misunderstanding may be related to the way the other two components, the Job Site and Hands-on, are handled. Job Site tasks are tested on-the-job and Hands-on tasks are tested at special test sites at which both training and testing take place. Preparation for these two components tends to be well organized and standardized from unit to unit and requires little initiative on the part of individual soldiers. The same passive approach soldiers are able to take to Hands-on and Job Site preparation seems to carry over to Skill Component preparation. SC performance might be improved if units would either organize and standardize written test preparation in the same way as the other two components are managed, or insure that individual soldiers' responsibilities are clearly understood.

CONCLUSIONS

Techniques for preparation for SC testing need to be re-examined. Guidelines for Hands-on and Job Site preparation follow naturally from explicit directions for testing and scoring set forth in SQT Notices. This is much less the case for written testing.

Written test preparation could be organized and standardized in the same way as are the other two components. One way to do this would be to administer written tests on SC tasks prior to official testing. If, however, soldiers are to assume individual responsibility for written test preparation, this requirement needs to be communicated clearly and training in individual study methods may have to be adopted.